PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference RE/VB60547	**FOR FURTHER ACTION**	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/EP 03/12793	International filing date *(day/month/year)* 13.11.2003	Priority date *(day/month/year)* 15.11.2002
International Patent Classification (IPC) or both national classification and IPC C07K14/18		
Applicant GLAXO GROUP LIMITED et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 9 sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 31 sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the opinion
   - II ☐ Priority
   - III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
13.05.2004	27.01.2005
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized Officer Irion, A Telephone No. +49 89 2399-8174

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP 03/12793

I. Basis of the report

1. With regard to the **elements** of the international application *(Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17))*:

Description, Pages

1-25, 28-33, 35-41	as originally filed
26, 27, 34	received on 01.04.2004 with letter of 31.03.2004

Claims, Numbers

1-20	filed with telefax on 03.11.2004

Drawings, Sheets

2/28-5/28, 8/28, 12/28-28/28	as originally filed
1/28, 6/28, 7/28, 9/28-11/28	received on 01.04.2004 with letter of 31.03.2004

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

   These elements were available or furnished to this Authority in the following language: , which is:

   ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).

   ☐ the language of publication of the international application (under Rule 48.3(b)).

   ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

   ☐ contained in the international application in written form.

   ☐ filed together with the international application in computer readable form.

   ☒ furnished subsequently to this Authority in written form.

   ☒ furnished subsequently to this Authority in computer readable form.

   ☒ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

   ☒ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

   ☐ the description, pages:

   ☐ the claims, Nos.:

   ☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)).

   (Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

   ☐ the entire international application,

   ☒ claims Nos. 18,19 (IA)

   because:

   ☒ the said international application, or the said claims Nos. 18,19 (IA) relate to the following subject matter which does not require an international preliminary examination (specify):

   see separate sheet

   ☐ the description, claims or drawings *(indicate particular elements below)* or said claims Nos. are so unclear that no meaningful opinion could be formed *(specify)*:

   ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

   ☐ no international search report has been established for the said claims Nos.

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

   ☐ the written form has not been furnished or does not comply with the Standard.

   ☐ the computer readable form has not been furnished or does not comply with the Standard.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-20
	No:	Claims	
Inventive step (IS)	Yes:	Claims	1, 6-12
	No:	Claims	2-5,13-20
Industrial applicability (IA)	Yes:	Claims	1-17,20
	No:	Claims	

2. Citations and explanations

see separate sheet

Item I

I.1 Sequence listing pages filed with letter of 31.03.2004 do not form part of the application (Rule 13ter.1(f) PCT).

I.2 The amendments filed with letter of 31.03.2004 and those filed with telefax of 03.11.2004 do not appear to introduce subject-matter which extends beyond the content of the application as filed (Article 34(2)(b) PCT.

Item III

III.1 With respect to claims 18 and 19

Claims 18 and 19 relate to subject-matter considered by this Authority to be covered by the provisions of Rule 67.1(iv) PCT. Consequently, no opinion will be formulated with respect to the industrial applicability of the subject-matter of these claims (Article 34(4)(I) PCT).

Item V

V.1 Reference is made to following documents

D1: WO0130812 (CHIRON CORPORATION) 03 May 2001 (2001-05-03)
D2: WO0138360 (CHIRON CORPORATION) 31 May 2001 (2001-05-31)
D3: WO9610997 (APOLLON, INC ET AL.) 18 April 1996 (1996-04-18)
D4: WO9747358 (CHIRON CORPORATION) 27 November 1997 (1997-11-27)
D5: J.P. MOORMAN ET AL.: 'The C-terminal region of hepatitis C core protein is required for Fas-ligand independent apoptosis in Jurkat cells by facilitating Fas oligomerization', VIROLOGY, 01 August 2003 (2003-08-01), vol. 312, pages 320-329

V.2 Novelty (Article 33(2) PCT)

V.2.1 With respect to claims 1-20

Document D1 describes plasmid DNA molecules encoding fusion proteins comprising (I) the full length Core protein or epitopes derived from the Core protein, and (ii) NS3, NS4a, NS4b, NS5a, and NS5b (p. 17 l. 22 - p. 18 l. 2). Said DNA molecules are used coupled to gold carriers for vaccination against HCV infection by a gene gun (p. 3 l.

23-27 and p. 22 l. 19-25).

Document D2 describes a vaccine against HCV comprising a fusion protein comprising truncated core (at amino acid 121), NS3, NS4a, NS4b, NS5a, NS5b (p. 3 l. 16 - p. 4 l. 9, p. 27 l. 1-19) or NS3-NS4b-NS5b combined with core (p. 26 l. 22 and p. 28 l. 11-14). Expression constructs comprising ΔNS3NS5 and either Core-121, Core-140, Core-150 or Core-173 within one expression cassette are described (p. 54 l. 27 - p. 56 l. 4). "Expression levels of the ΔNS3NS5-Core-173 construct were much less than that of the ΔNS3NS5-Core-121 construct" and D2 states that "there is a correlation of protein expression levels and the length of HCV core" (p. 56 l. 16-18). Furthermore, the constructs comprising Core-140 or Core-150 were expressed at a similar level as the ΔNS3NS5-Core-173 construct (p. 56 l. 20-22). The NS3 protein is encoded by a nucleic acid sequence having an N-terminal deletion to remove the catalytic domain. Said polypeptide comprises a deletion in, or mutation of, the NS3 protease active site region to render the protease non-functional (p. 10 l. 27 - p. 11 l. 7). The polypeptide comprising the proteins before-mentioned and the DNA polynucleotide molecule encoding said polypeptide are described (p. 4 l. 24-31). Gold particles coated with the DNA molecule used for vaccination by gene gun are described (p. 44 l. 17-22). Said DNA may be comprised in a plasmid. A method of eliciting an immune response against HCV using the polynucleotide mentioned above is described (p. 6 l. 23-25).

Thus, none of the documents cited in the international search report disclose the subject-matter as defined in claims 1-20, i.e. the HCV proteins are encoded by the polynucleotide vaccine in more than one expression cassette. Therefore, said claims are considered novel in the sense of Article 33(2) PCT.

V.3 Inventive step (Article 33(3) PCT)

V.3.1 With respect to claims 1 and 6-12

The subject-matter of claims 1 and 6-12 differs from the closest prior art document D2 in that the expression cassette encoding the Core protein is downstream of the expression cassette which encodes at least one of the other HCV proteins. The technical problem to be solved may be regarded as providing an alternative HCV vaccine. None of the documents cited in the international search report suggests that the position of the polynucleotide encoding the Core protein downstream of the other expression cassette would result in an increased expression level of the other HCV proteins, for which experimental evidence is given in the Example 6 of the application.

Therefore, the subject-matter of claims 1 and 6-12 is considered inventive in the sense of Article 33(3) PCT.

V.3.2 With respect to claim 17

The subject-matter of claim 17 differs from the closest prior art document D2 in that the specific Core truncates are disclosed, i.e. Core-151, Core-165, Core-171. Document D2 shows that fusion proteins comprising Core-173, Core-140 or Core-150 are expressed at low levels (p. 56 l. 16-22). Therefore, none of the documents cited in the international search report suggests that said truncates would result in an increased expression level of the other HCV protein. The present application gives experimental evidence in the Example 7 that Core truncates Core-151 and Core-171 show the alleged effect. The subject-matter of claim 17, limited to Core-151 and Core-171, would be considered inventive in the sense of Article 33(3) PCT. However, for the Core-165 no experimental data are given. Therefore, it is not clear whether said truncate solves the technical problem posed. Therefore, the subject-matter of claim 17 is not considered inventive in the sense of Article 33(3) PCT.

V.3.3 With respect to claims 2-5, 13-16, and 18-20

The subject-matter of claim 2 differs from the closest prior art document D2 in that the core protein used is encoded in a separate expression cassette. The problem to be solved by the subject-matter of claims 2-5,13-16, and 18-20 may be regarded as to provide an alternative HCV vaccine. The solution provided in claims 2-5, 13-16, and 18-20 resides in the use of more than one expression cassettes. No surprising effect is shown in the application of the use of more than one expression cassettes instead of only one comprising the polynucleotides encoding the Core fragments as defined in claims 2-5.

Furthermore, the subject-matter of claim 16 differs from the closest prior art document D2 in that the HCV proteins used for vaccination are not codon optimised. The technical problem to be solved may be regarded as the provision of a HCV vaccine which is expressed efficiently in the human organism. The person skilled in the art is aware of the fact, that codon pairings are highly nonrandom and differ from organism to organism, resulting in a low translational efficiency. The solution provided in claim 16 resides in the use of codon optimised polynucleotides for the expression of the HCV antigens. However, the skilled person would combine the teaching of document D4, which describes the production of codon optimised expression of HCV proteins (p. 5 l. 29 - p. 10 l. 8, p. 18 l. 8-21, Figures 12 and 13), with D2 to solve the

problem of low translational efficiency.

Therefore, the subject-matter of claims 2-5, 13-16, and 18-20 is not considered inventive in the sense of Article 33(3) PCT.

V.4 Industrial applicability (Article 33(4) PCT)

V.4.1 With respect to claims 1-17 and 20

The subject-matter of claims 1-17 and 20 appears to be susceptible of industrial application.

V.4.2 With respect to claims 18 and 19

The subject-matter of claims 18 and 19 is considered to be a method of treatment by therapy of the human or animal body.

For the assessment of the present claims 18 and 19 on the question whether they are industrially applicable, no unified criteria exist in the PCT Contracting States. The patentability can also be dependent upon the formulation of the claims. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but may allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

V.5 Remark concerning document D5

The examination report has been based on an assumed valid priority for the present application. Should the priority of the present application not be valid, the above cited document D5 would be relevant with respect to novelty and inventive step (Article 33(2) and (3) PCT).

V.6 Further remarks

V.6.1 With respect to claims 2, 13-16, and 18-20

The subject-matter of claims 2, 13-16, and 18-20 does not meet the requirements of Article 6 PCT in that the matter for which protection is sought is not clearly defined. The claims attempt to define the subject-matter in terms of the result to be achieved which merely amounts to a statement of the underlying problem, i.e. the mutation of the core protein sequence such that the negative effect of expression of the Core protein upon the expression of the other HCV protein(s) is reduced. The technical

features necessary for achieving this result should be added.

V.6.2 With respect to claim 3

The subject-matter of claim 3 does not meet the requirements of Article 6 PCT in that the matter for which protection is sought is not clearly defined. The claim attempts to define the subject-matter in terms of the result to be achieved which merely amounts to a statement of the underlying problem, i.e. the truncation from the C-terminal end in a sufficient amount to reduce the inhibitory effect of Core upon the expression of other HCV proteins. The technical features necessary for achieving this result should be added.

HCV Core
Forward primer (SEQ ID NO. 1)
5'-GAATTC**GCGGCCGCCATG**AGCACCAACCCCAAGCCCCAGCGCAAGACCAAGCGGAACACC-3'
NotI translation start codon

Reverse primer (SEQ ID NO. 2)
5'-GAATTC**GGATCCTCA**TGCGCTAGCGGGGATGGTGAGGCAGCTCAGCAGCGCCAGCAGGA-3'
BamHI Stop codon

HCV NS3
Forward primer (SEQ ID NO. 3)
5'-GAATTC**GCGGCCGCCATG**GCCCCCATCACCGCCTACAGCCAGCAGACCCGGGGAC-3'
NotI translation start codon

Reverse primer (SEQ ID NO. 4)
5'-GAATTC**GGATCCTCA**GGTGACCACCTCCAGGTCAGCGGACATGCACGCCATGATG-3'
BamHI Stop codon

HCV NS4B
Forward primer (SEQ ID NO. 5)
5'-GAATTC**GCGGCCGCCATG**TTTTGGGCCAAGCATATGTGGAACTTCA-3'
NotI translation start codon

Reverse primer (SEQ ID NO. 6)
5'-GAATTC**GGATCCTCA**GCAAGGGGTGGAGCAGTCCTCGTTGATCCAC-3'
BamHI Stop codon

HCV NS5B
Forward primer (SEQ ID NO. 7)
5'-GAATTC**GCGGCCGCCATG**TCCATGTCCTACACCTGGACCGGCGCCCTGA-3'
NotI translation start codon
Reverse primer (SEQ ID NO. 8)
5'-GAATTC**GGATCCTCA**GCGGTTGGGCAGCAGGTAGATGCCGACTCCGACG-3'
BamHI Stop codon

All polynucleotides, encoding single antigens, were cloned into mammalian expression vector p7313ie via Not I and BamHI unique cloning sites (see figure 7).

The polyproteins that were encoded were as follows (including mutations and codon optimisations):

HCV Core translation (SEQ ID NO. 9):

MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERS
QPRGRRQPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDP

RRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGN
LPGCSFSIFLLALLSCLTIPASA

HCV NS3 translation (SEQ ID NO. 10):

MAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCINGVCWTVY
HGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSMTPCTCGSSDLYLVTRHA
DVIPVRRRGDSRGSLLSPRPVSYLKGSVGGPLLCPSGHVVGIFRAAVCTRGVAKAVD
FIPVESMETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQGYKV
LVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYGKFLADGGCSGGA
YDIIICQECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSN
NGEIPFYGKAIPIEAIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPT
SGDVVVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRS
QRRGRTGRGRSGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAY
LNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQ
APPPSWDQMWKCLIRLKPTLHGPTPLLYRLGAVQNEVTLTHPITKYIMACMSADLEV
VT

HCV NS4B translation (SEQ ID NO. 11):

MFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQNTLLFNILGGWV
AAQLAPPSAASAFVGAGIAGAAVGSIGLGKVLVDILAGYGAGVAGALVAFKVMSGE
VPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNH
VSPTHYVPESDAAARVTQILSSLTITQLLKRLHQWINEDCSTPC

HCV NS5B translation (SEQ ID NO. 12):

MSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDR
LQVLDDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLS
SRAVNHIRSVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVC
EKMALYDVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYGTRCFG
STVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASG
VLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAF
TEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAA
WETARHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYS
IEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLL
SQGGRAATCGRYLFNWAVRTKLKLTPIPAASQLDLSGWFVAGYSGGDIYHSLSRAR
PRWFPLCLLLLSVGVGIYLLPNR

Example 3, ***Immune response assays***

Table 3 *Frequency of NS4B CD4 or CD8 specific T cell producing IFNγ following immunisation with HCV polyproteins.*

Plasmid	nil	NS4B protein	NS4B CD4 peptide	NS4B CD8 peptide
NS4B	0.05	0.17	0.18	2.04
HCV500	0.09	0.09	0.1	0.6
HCV510	0.05	0.09	0.09	0.34
HCV520	0.06	0.08	0.05	0.33
HCV530	0.1	0.17	0.1	0.37
HCV501	0.04	0.09	0.06	0.13

IFNγ specific T cell responses were detected following of stimulation of splenocytes in presence or absence of antigen for 6 hours, in presence of Brefeldin A for last 4hours. IFNg was detected by gating on CD4 or CD8 T cells and staining with IFNγ FITC.

The peptides used have following sequence:

Protein	Peptides
NS3	(C57Bl) CD4 PRFGKAIPIEAIKGG (SEQ ID NO. 13) CD8 YRLGAVQNEVILTHP (SEQ ID NO. 14)
NS5	(C57BL/6). CD4 SMSYTWTGALITPCA (SEQ ID NO. 15) CD8 AAALRAFTEAMTRYS (SEQ ID NO. 16)
NS4B	(Balb/c) CD4 IQYLAGLSTLPGNPA (SEQ ID NO. 17) CD8 FWAKHMWNFISGIWY (SEQ ID NO. 18)

Recognition of endogenously processed antigen

In order to determine if PMID immunisation with the HCV polyproteins induced a response that could recognise endogenously processed antigen, targets cells infected with Vaccinia recombinant virus expressing NS3-5 were used as stimulators in the ELISPOT

Claims

1. A polynucleotide vaccine comprising a polynucleotide sequence that encodes the HCV Core protein and a polynucleotide sequence that encodes at least one other HCV protein, wherein the vaccine causes expression of the proteins within the same cell wherein the Core protein and the at least one other HCV protein are encoded in more than one expression cassette characterised in that the expression cassette encoding the Core protein is in a cis location downstream of the expression cassette which encodes at least one of the other HCV proteins.

2. A polynucleotide vaccine comprising a polynucleotide sequence that encodes the HCV Core protein and a polynucleotide sequence that encodes at least one other HCV protein, wherein the vaccine causes expression of the proteins within the same cell and the sequence of the polynucleotide sequence encoding the core protein has been mutated such that the negative effect of expression of the Core protein upon the expression of the said at least one other HCV protein is reduced, wherein the HCV proteins are encoded by the polynucleotide vaccine in more than one expression cassettes.

3. A polynucleotide vaccine as claimed in claim 1 or 2, wherein polynucleotide encodes a core protein that is truncated from the carboxy terminal end in a sufficient amount to reduce the inhibitory effect of Core upon the expression of other HCV proteins.

4. A polynucleotide vaccine as claimed in claim 3 wherein the polynucleotide encodes the mature form of HCV core protein after the second naturally occurring cleavage during normal HCV infection.

5. A polynucleotide vaccine as claimed in 3 wherein the truncated core protein has a deletion of at least the C-terminal 10 amino acids.

6. A polynucleotide vaccine as claimed in claim 3 wherein the truncated core protein consists of the Core 1-151 sequence.

7. A polynucleotide vaccine as claimed in claim 3 wherein the truncated core protein consists of the Core 1-165 sequence.

8. A polynucleotide vaccine as claimed in claim 1 or claim 2 wherein the expression cassette encoding the Core protein is downstream of an expression cassette that encodes the NS5B protein.

9. A polynucleotide vaccine as claimed in claim 8 wherein the expression cassette encoding the Core protein encodes for Core protein in fusion with the HCV NS3 protein.

10. An HCV vaccine as claimed in claim 8, wherein one expression cassette encodes the double fusion protein NS3-Core and the other encoding a NS4B-NS5B double fusion protein.

11. An HCV vaccine as claimed in claim 10 wherein the Core element of the NS3-Core double fusion protein is selected from the group consisting of Core 1-171, Core 1-165 and Core 1-151.

12. An HCV vaccine as claimed in claim 11, wherein the Core element of the NS3-Core double fusion protein is Core 1-165.

13. A polynucleotide vaccine as claimed in claim 1 or claim 2, wherein the at least one other HCV protein comprises the HCV proteins: NS3, NS4B and NS5B.

14. A polynucleotide vaccine as claimed in claim 13, wherein the polynucleotide encodes no other HCV protein.

15. A polynucleotide vaccine as claimed in any one of claims 1 to 14 wherein the polynucleotide sequence is in the form of a plasmid.

16. A polynucleotide vaccine as claimed in any one of claims 1 to 14 wherein the polynucleotides are codon optimised for expression in mammalian cells.

17. A polynucleotide vaccine comprising a polynucleotide sequence that encodes the HCV Core protein and a polynucleotide sequence that encodes at least one other HCV

protein, wherein the vaccine causes expression of the proteins within the same cell and the sequence of the polynucleotide sequence encoding the core protein has been mutated or positioned relative to the polynucleotide sequence encoding the at least one other HCV protein such that the negative effect of expression of the Core protein upon the expression of the said at least one other HCV protein is reduced, characterised in that the Core protein encoded by the polynucleotide vaccine consists of one of the following group of sequences: Core 1-151, Core 1-165 and Core 1-171.

18. A method of preventing or treating an HCV infection in a mammal comprising administering a vaccine as claimed in any one of claims 1 to 17 to a mammal.

19. A method of vaccination of an individual comprising taking a polynucleotide vaccine as claimed in any one of claims 1 to 17, coating the polynucleotide onto gold beads and delivering the gold beads into the skin.

20. Use of a polynucleotide vaccine as claimed in any one of claims 1 to 17 in the manufacture of a medicament for the treatment of HCV.

Figure 1, HCV J4L6 genome wild-type cDNA sequence, reference accession number AF054247 (SEQ ID NO. 19),

```
   1 gccagccccc tgatgggggc gacactccac catgaatcac tcccctgtga ggaactactg
  61 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac
 121 ccccctcccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag
 181 gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc
 241 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg
 301 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac
 361 ctcaaagaaa aaccaaacgt aacaccaacc gccgcccaca ggacgtcaag ttcccgggcg
 421 gtggtcagat cgttggtgga gtttacctgt tgccgcgcag gggccccagg ttgggtgtgc
 481 gcgcgactag gaaggcttcc gagcggtcgc aacctcgtgg aaggcgacaa cctatcccaa
 541 aggctcgccg acccgagggc agggcctggg ctcagcccgg gtacccttgg ccccctctatg
 601 gcaatgaggg cctggggtgg gcaggatggc tcctgtcacc ccgcggctcc cggcctagtt
 661 ggggccccac ggacccccgg cgtaggtcgc gtaacttggg taaggtcatc gataccctta
 721 catgcggctt cgccgatctc atggggtaca ttccgctcgt cggcgccccc ctaggggcg
 781 ctgccagggc cttggcacac ggtgtccggg ttctggagga cggcgtgaac tatgcaacag
 841 ggaacttgcc cggttgctct ttctctatct tcctcttggc tctgctgtcc tgtttgacca
 901 tcccagcttc cgcttatgaa gtgcgcaacg tgtccgggat ataccatgtc acgaacgact
 961 gctccaactc aagcattgtg tatgaggcag cggacgtgat catgcatact cccgggtgcg
1021 tgccctgtgt tcaggagggt aacagctccc gttgctgggt agcgctcact cccacgctcg
1081 cggccaggaa tgccagcgtc cccactacga caatacgacg ccacgtcgac ttgctcgttg
1141 ggacggctgc tttctgctcc gctatgtacg tgggggatct ctgcggatct attttcctcg
1201 tctcccagct gttcaccttc tcgcctcgcc ggcatgagac agtgcaggac tgcaactgct
1261 caatctatcc cggccatgta tcaggtcacc gcatggcttg ggatatgatg atgaactggt
1321 cacctacaac agccctagtg gtgtcgcagt tgctccggat cccacaagct gtcgtggaca
1381 tggtggcggg ggcccactgg ggagtcctgg cgggccttgc ctactattcc atggtaggga
1441 actgggctaa ggttctgatt gtggcgctac tctttgccgg cgttgacggg gagacccaca
1501 cgacggggag ggtggccggc cacaccacct ccgggttcac gtccctttтc tcatctgggg
1561 cgtctcagaa aatccagctt gtgaatacca acggcagctg gcacatcaac aggactgccc
1621 taaattgcaa tgactccctc caaactgggt tctttgccgc gctgttttac gcacacaagt
1681 tcaactcgtc cgggtgcccg gagcgcatgg ccagctgccg ccccattgac tggttcgccc
1741 aggggtgggg ccccatcacc tatactaagc ctaacagctc ggatcagagg ccttattgct
1801 ggcattacgc gcctcgaccg tgtggtgtcg tacccgcgtc gcaggtgtgt ggtccagtgt
1861 attgtttcac cccaagccct gttgtggtgg ggaccaccga tcgttccggt gtccctacgt
1921 atagctgggg ggagaatgag acagacgtga tgctcctcaa caacacgcgt ccgccacaag
1981 gcaactggtt cggctgtaca tggatgaata gtactgggtt cactaagacg tgcggaggtc
2041 ccccgtgtaa catcgggggg gtcggtaacc gcaccttgat ctgccccacg gactgcttcc
2101 ggaagcaccc cgaggctact tacacaaaat gtggctcggg gccctggttg acacctaggt
```

Figure 2, codon optimised HCV Core polynucleotide (SEQ ID NO. 20)

ATGAGCACCAACCCCAAGCCCCAGCGCAAGACCAAGCGGAACACCAACCGGAGACCCCAGGA
CGTCAAGTTCCCAGGAGGAGGCCAGATCGTGGGCGGCGTGTACCTGCTGCCCCGCCGGGGGC
CCCGGCTGGGCGTGCGCGCCACCCGCAAGACCAGCGAGCGCTCCCAGCCAAGAGGCAGACGC
CAGCCGATCCCGAAGGCCCGCCGCCCTGAGGGCCGGGCTTGGGCCCAGCCAGGCTACCCCTG
GCCCCTGTATGGCAACGAGGGCCTGGGATGGGCTGGGTGGCTCCTCAGCCCCCGGGGGTCTA
GGCCCAGTTGGGGACCGACCGACCCCCGCAGGCGCAGCCGCAACCTGGGAAAGGTGATCGAC
ACGCTCACCTGCGGCTTCGCCGACTTGATGGGATACATCCCTCTGGTGGGGGCCCCTCTGGG
CGGAGCCGCGCGCGCCCTGGCTCACGGGGTCCGGGTGCTCGAGGACGGGGTGAACTACGCCA
CCGGGAACCTGCCCGGCTGCAGCTTCTCCATCTTCCTGCTGGCGCTGCTGAGCTGCCTCACC
ATCCCCGCTAGCGCATGA

6/28

Figure 3, Codon optimised HCV NS3 polynucleotide (SEQ ID NO. 21)

ATGGCCCCCATCACCGCCTACAGCCAGCAGACCCGGGGACTGCTCGGCTGCATCATCACCTC
TCTGACAGGCCGGGATAAGAACCAGGTGGAGGGCGAGGTGCAGGTCGTCTCGACCGCTACCC
AAAGCTTCCTGGCCACCTGTATCAACGGAGTCTGCTGGACGGTGTACCATGGCGCCGGCAGC
AAGACCCTCGCCGGGCCTAAGGGCCCCATCACCCAGATGTACACCAACGTGGACCAGGACCT
GGTGGGCTGGCAGGCGCCCCCCGGGGCGAGGAGTATGACCCCATGCACCTGCGGGAGCTCTG
ACCTGTATCTGGTGACCAGACATGCCGATGTCATCCCGGTGAGGCGTCGCGGGGACAGTAGA
GGGAGCCTGCTGAGCCCCCGCCCCGTCAGCTACCTGAAGGGGTCCGTGGGCGGCCCCCTGCT
GTGCCCCTCTGGCCACGTGGTCGGCATCTTCAGGGCCGCCGTGTGCACGCGCGGCGTGGCCA
AGGCCGTGGACTTTATCCCCGTGGAGAGCATGGAGACCACCATGCGCTCCCCCGTGTTCACC
GACAACAGCAGCCCCCCCGCCGTGCCTCAGACCTTCCAGGTCGCCCACCTCCATGCTCCGAC
GGGCTCCGGGAAGTCCACGAAGGTGCCCGCCGCGTACGCGGCCCAGGGATACAAGGTGCTGG
TCCTCAACCCTAGCGTGGCTGCCACACTCGGGTTTGGAGCGTACATGAGCAAGGCGCACGGC
ATCGACCCCAACATCAGAACTGGCGTCCGGACCATCACAACCGGCGCTCCCATCACTTACTC
TACCTACGGCAAGTTCCTGGCTGATGGGGGGTGTAGTGGGGGCGCGTACGATATTATCATCT
GCCAGGAGTGCCACTCTACCGACAGCACCACAATCCTGGGCATCGGCACCGTCCTCGACCAG
GCTGAGACAGCGGGCGCCCGCCTGGTGGTGCTGGCCACGGCCACTCCCCCCGGCTCCGTCAC
GGTGCCCCACCCCAATATCGAGGAGGTGGCCCTGAGCAACAACGGCGAGATCCCATTCTACG
GCAAGGCTATCCCGATCGAGGCGATTAAGGGAGGCAGACATCTGATCTTCTGCCACAGCAAG
AAGAAGTGCGACGAGCTCGCCGCCAAGCTGAGCGGCCTCGGACTCAACGCCGTGGCTTACTA
CAGGGGACTGGACGTGTCCGTGATCCCGACCAGCGGAGACGTGGTGGTCGTGGCCACCGACG
CCCTGATGACCGGCTTCACCGGAGACTTCGACAGCGTCATCGACTGCAACACCTGCGTGACC
CAGACCGTGGACTTCAGCCTGGACCCCACCTTCACCATCGAGACCACCACAGTGCCCCAGGA
CGCCGTGTCCCGCAGCCAGCGCCGGGGCCGGACCGGCCGCGGCCGGAGTGGCATCTATAGGT
TCGTGACCCCGGGCGAGCGCCCCAGCGGCATGTTCGATAGTTCCGTGCTGTGCGAGTGCTAC
GACGCCGGATGCGCGTGGTACGAGCTGACCCCGGCGGAGACCTCTGTCCGCCTGAGGGCTTA
CTTGAATACCCCGGGCCTGCCCGTGTGCCAGGATCATCTCGAGTTCTGGGAATCCGTCTTCA
CCGGCCTGACACACATCGACGCCCATTTCTTGTCCCAAACCAAGCAGGCTGGCGACAATTTC
CCGTATCTGGTCGCGTACCAGGCCACGGTGTGCGCGCGTGCGCAGGCTCCCCCCCCTAGCTG
GGATCAGATGTGGAAGTGCCTGATCCGCCTGAAGCCCACCCTGCATGGGCCCACCCCCCTGC
TGTACCGCCTGGGCGCGGTGCAGAACGAAGTCACCTTGACCCACCCCATCACCAAGTACATC

Figure 4, codon optimised HCV NS4B polynucleotide (SEQ ID NO. 22)

ATGTTTTGGGCCAAGCATATGTGGAACTTCATCAGCGGCATCCAGTACCTCGCCGGGCTGAG
CACCCTCCCGGGCAACCCCGCGATCGCAAGCCTGATGGCGTTCACAGCGAGCATCACCTCCC
CCCTGACTACCCAGAACACACTGCTGTTCAACATCCTGGGGGGCTGGGTCGCCGCTCAGCTG
GCCCCTCCTTCCGCCGCCAGCGCCTTTGTGGGGGCGGGAATCGCCGGGGCCGCCGTCGGCTC
CATCGGACTGGGCAAGGTGCTGGTCGACATCCTGGCGGGCTACGGCGCGGGAGTCGCCGGAG
CCCTGGTGGCCTTCAAGGTGATGAGCGGAGAGGTGCCAAGCACTGAGGACCTGGTGAACCTG
CTGCCGGCGATCCTGAGCCCGGGCGCCCTGGTGGTGGGCGTGGTGTGTGCTGCCATCCTCAG
GCGCCACGTGGGCCCGGGCGAGGGAGCCGTGCAGTGGATGAACCGCCTGATCGCCTTTGCCT
CCCGCGGCAACCACGTCAGCCCTACACATTACGTGCCCGAGAGCGATGCCGCCGCCCGCGTG
ACCCAGATCCTGAGCTCCCTGACCATCACCCAGCTGCTCAAGAGGCTGCACCAGTGGATCAA
CGAGGACTGCTCCACCCCTTGCTGA

9/28

Figure 5, codon optimised HCV NS5B polynucleotide (SEQ ID NO. 23)

ATGTCCATGTCCTACACCTGGACCGGCGCCCTGATCACCCCCTGCGCCGCCGAGGAGAGCAA
GCTCCCGATTAACCCCCTGTCCAACTCTCTGCTCCGCCATCACAACATGGTGTATGCCACCA
CCTCCCGCTCTGCGAGCCTCCGCCAGAAGAAGGTGACGTTCGACAGACTGCAGGTGCTGGAC
GACCATTACAGGGACGTGCTGAAGGAAATGAAGGCCAAGGCTAGCACCGTGAAGGCCAAGCT
GCTCAGCATTGAGGAGGCTTGCAAGCTGACCCCCCCCCACAGTGCTAAATCCAAGTTCGGCT
ACGGCGCCAAGGACGTGAGGAACCTGTCCTCGCGCGCTGTGAACCATATCCGCAGCGTGTGG
GAGGACCTGCTCGAGGACACCGAGACCCCCATCGACACAACCATCATGGCCAAGTCCGAGGT
GTTCTGCGTGCAGCCGGAGAAAGGAGGCCGCAAGCCAGCCCGCCTGATCGTCTTCCCCGACC
TGGGCGTGAGAGTCTGCGAGAAGATGGCCCTCTACGACGTGGTGTCCACCCTGCCGCAGGCC
GTGATGGGGAGTTCCTACGGCTTCCAGTACAGCCCGAAGCAGAGGGTGGAGTTCCTGGTGAA
CACGTGGAAGTCTAAGAAATGCCCCATGGGGTTCAGTTACGGAACAAGGTGCTTCGGGAGTA
CTGTGACCGAATCCGATATCCGCGTGGAGGAGAGCATCTACCAGTGTTGTGACCTCGCCCCC
GAGGCGAGACAGGCCATCCGCTCCCTGACCGAGAGGCTGTATATCGGCGGCCCACTGACCAA
CAGCAAGGGGCAGAACTGCGGCTATCGCCGTTGTCGGGCCTCCGGGGTGCTCACCACCTCTT
GCGGGAACACCCTCACCTGCTACCTCAAGGCGACCGCTGCCTGCAGAGCCGCGAAGCTGCAG
GACTGCACCATGCTCGTGAACGGCGACGATCTGGTGGTGATCTGTGAGTCCGCGGGCACGCA
GGAGGACGCGGCGGCCCTGCGGGCGTTCACAGAGGCCATGACACGCTACAGTGCCCCCCCCG
GCGACCCCCCCCAGCCCGAATACGATCTGGAGCTCATCACTAGTTGCAGCTCGAACGTGTCT
GTGGCCCATGACGCTTCTGGCAAACGGGTGTATTATCTGACGCGCGATCCCACCACCCCCCT
CGCCAGAGCCGCGTGGGAGACAGCTCGGCACACCCCTGTGAACTCTTGGCTGGGCAACATCA
TCATGTACGCCCCTACCCTGTGGGCTCGCATGATCCTGATGACCCACTTCTTCAGTATCCTC
CTCGCTCAGGAGCAGCTGGAGAAGGCGCTCGACTGCCAGATCTACGGCGCCTGCTATAGTAT
CGAGCCTCTCGACCTGCCCCAGATCATCGAGAGACTGCATGGGCTCAGCGCCTTCTCCCTCC
ATAGTTACTCTCCTGGAGAAATTAACCGGGTGGCGAGCTGTCTGCGGAAGCTCGGCGTCCCC
CCTCTGCGCGTTTGGCGGCATCGCGCCAGGAGTGTGAGGGCCAAGCTGCTGAGCCAGGGCGG
AAGGGCCGCCACCTGCGGCCGGTATCTCTTCAACTGGGCCGTGCGCACCAAGCTCAAGCTCA
CCCCCATCCCTGCCGCCAGTCAGCTGGATCTCAGTGGGTGGTTCGTGGCCGGCTATTCTGGC
GGCGACATCTACCACTCCCTCAGCAGGGCGCGCCCCCGCTGGTTCCCCCTGTGCCTGCTGCT
CCTGAGCGTCGGAGTCGGCATCTACCTGCTGCCCAACCGCTGA

Figure 6, *Translation of HCV J4L6 genome (wild-type sequence)* (SEQ ID NO. 24)

```
   1 MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR KASERSQPRG
  61 RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP RGSRPSWGPT DPRRRSRNLG
 121 KVIDTLTCGF ADLMGYIPLV GAPLGGAARA LAHGVRVLED GVNYATGNLP GCSFSIFLLA
 181 LLSCLTIPAS AYEVRNVSGI YHVTNDCSNS SIVYEAADVI MHTPGCVPCV QEGNSSRCWV
 241 ALTPTLAARN ASVPTTTIRR HVDLLVGTAA FCSAMYVGDL CGSIFLVSQL FTFSPRRHET
 301 VQDCNCSIYP GHVSGHRMAW DMMMNWSPTT ALVVSQLLRI PQAVVDMVAG AHWGVLAGLA
 361 YYSMVGNWAK VLIVALLFAG VDGETHTTGR VAGHTTSGFT SLFSSGASQK IQLVNTNGSW
 421 HINRTALNCN DSLQTGFFAA LFYAHKFNSS GCPERMASCR PIDWFAQGWG PITYTKPNSS
 481 DQRPYCWHYA PRPCGVVPAS QVCGPVYCFT PSPVVVGTTD RSGVPTYSWG ENETDVMLLN
 541 NTRPPQGNWF GCTWMNSTGF TKTCGGPPCN IGGVGNRTLI CPTDCFRKHP EATYTKCGSG
 601 PWLTPRCLVD YPYRLWHYPC TLNFSIFKVR MYVGGVEHRL NAACNWTRGE RCNLEDRDRS
 661 ELSPLLLSTT EWQILPCAFT TLPALSTGLI HLHQNIVDVQ YLYGVGSAFV SFAIKWEYIL
 721 LLFLLLADAR VCACLWMMLL IAQAEAALEN LVVLNAASVA GAHGILSFLV FFCAAWYIKG
 781 RLAPGAAYAF YGVWPLLLLL LALPPRAYAL DREMAASCGG AVLVGLVFLT LSPYYKVFLT
 841 RLIWWLQYFI TRAEAHMQVW VPPLNVRGGR DAIILLTCAV HPELIFDITK LLLAILGPLM
 901 VLQAGITRVP YFVRAQGLIR ACMLVRKVAG GHYVQMVFMK LGALTGTYVY NHLTPLRDWA
 961 HAGLRDLAVA VEPVVFSAME TKVITWGADT AACGDIILGL PVSARRGKEI FLGPADSLEG
1021 QGWRLLAPIT AYSQQTRGVL GCIITSLTGR DKNQVEGEVQ VVSTATQSFL ATCINGVCWT
1081 VYHGAGSKTL AGPKGPITQM YTNVDLDLVG WQAPPGARSM TPCSCGSSDL YLVTRHADVI
1141 PVRRRGDSRG SLLSPRPVSY LKGSSGGPLL CPSGHVVGVF RAAVCTRGVA KAVDFIPVES
1201 METTMRSPVF TDNSTPPAVP QTFQVAHLHA PTGSGKSTKV PAAYAAQGYK VLVLNPSVAA
1261 TLGFGAYMSK AHGIDPNIRT GVRTITTGGS ITYSTYGKFL ADGGCSGGAY DIIICDECHS
1321 TDSTTILGIG TVLDQAETAG ARLVVLATAT PPGSVTVPHP NIEEIGLSNN GEIPFYGKAI
1381 PIEAIKGGRH LIFCHSKKKC DELAAKLTGL GLNAVAYYRG LDVSVIPPIG DVVVVATDAL
1441 MTGFTGDFDS VIDCNTCVTQ TVDFSLDPTF TIETTTVPQD AVSRSQRRGR TGRGRSGIYR
1501 FVTPGERPSG MFDSSVLCEC YDAGCAWYEL TPAETSVRLR AYLNTPGLPV CQDHLEFWES
1561 VFTGLTHIDA HFLSQTKQAG DNFPYLVAYQ ATVCARAQAP PPSWDQMWKC LIRLKPTLHG
1621 PTPLLYRLGA VQNEVILTHP ITKYIMACMS ADLEVVTSTW VLVGGVLAAL AAYCLTTGSV
1681 VIVGRIILSG KPAVVPDREV LYQEFDEMEE CASQLPYIEQ GMQLAEQFKQ KALGLLQTAT
1741 KQAEAAAPVV ESKWRALETF WAKHMWNFIS GIQYLAGLST LPGNPAIASL MAFTASITSP
1801 LTTQNTLLFN ILGGWVAAQL APPSAASAFV GAGIAGAAVG SIGLGKVLVD ILAGYGAGVA
1861 GALVAFKVMS GEVPSTEDLV NLLPAILSPG ALVVGVVCAA ILRRHVGPGE GAVQWMNRLI
1921 AFASRGNHVS PTHYVPESDA AARVTQILSS LTITQLLKRL HQWINEDCST PCSGSWLRDV
1981 WDWICTVLTD FKTWLQSKLL PRLPGVPFLS CQRGYKGVWR GDGIMQTTCP CGAQIAGHVK
2041 NGSMRIVGPR TCSNTWHGTF PINAYTTGPC TPSPAPNYSR ALWRVAAEEY VEVTRVGDFH
2101 YVTGMTTDNV KCPCQVPAPE FFTEVDGVRL HRYAPACKPL LREDVTFQVG LNQYLVGSQL
```

11/28

SEQUENCE LISTING

<110> Glaxo Group Ltd

<120> Vaccine

<130> VB60547

<140> PCT/EP03/12793
<141> 2003-11-13

<160> 24

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 60
<212> DNA
<213> Hepatitis C virus

<400> 1
gaattcgcgg ccgccatgag caccaacccc aagccccagc gcaagaccaa gcggaacacc 60

<210> 2
<211> 59
<212> DNA
<213> Hepatitis C virus

<400> 2
gaattcggat cctcatgcgc tagcggggat ggtgaggcag ctcagcagcg ccagcagga 59

<210> 3
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 3
gaattcgcgg ccgccatggc ccccatcacc gcctacagcc agcagacccg gggac 55

<210> 4
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 4
gaattcggat cctcaggtga ccacctccag gtcagcggac atgcacgcca tgatg 55

<210> 5
<211> 46
<212> DNA
<213> Hepatitis C virus

<400> 5
gaattcgcgg ccgccatgtt ttgggccaag catatgtgga acttca 46

<210> 6
<211> 46
<212> DNA
<213> Hepatitis C virus

<400> 6
gaattcggat cctcagcaag gggtggagca gtcctcgttg atccac 46

<210> 7
<211> 49
<212> DNA
<213> Hepatitis C virus

<400> 7
gaattcgcgg ccgccatgtc catgtcctac acctggaccg gcgccctga 49

<210> 8
<211> 49
<212> DNA
<213> Hepatitis C virus

<400> 8
gaattcggat cctcagcggt tgggcagcag gtagatgccg actccgacg 49

<210> 9
<211> 191
<212> PRT
<213> Hepatitis C virus

<400> 9

Met	Ser	Thr	Asn	Pro	Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	Asn	Thr	Asn
1				5					10					15	
Arg	Arg	Pro	Gln	Asp	Val	Lys	Phe	Pro	Gly	Gly	Gly	Gln	Ile	Val	Gly
			20					25					30		
Gly	Val	Tyr	Leu	Leu	Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg	Ala
		35					40					45			
Thr	Arg	Lys	Thr	Ser	Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln	Pro
	50					55					60				
Ile	Pro	Lys	Ala	Arg	Arg	Pro	Glu	Gly	Arg	Ala	Trp	Ala	Gln	Pro	Gly
65					70					75					80
Tyr	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Leu	Gly	Trp	Ala	Gly	Trp
				85					90					95	
Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	Pro
			100					105					110		
Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Cys
		115					120					125			
Gly	Phe	Ala	Asp	Leu	Met	Gly	Tyr	Ile	Pro	Leu	Val	Gly	Ala	Pro	Leu
	130					135					140				
Gly	Gly	Ala	Ala	Arg	Ala	Leu	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp
145					150					155					160
Gly	Val	Asn	Tyr	Ala	Thr	Gly	Asn	Leu	Pro	Gly	Cys	Ser	Phe	Ser	Ile
				165					170					175	
Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Leu	Thr	Ile	Pro	Ala	Ser	Ala	
			180					185					190		

<210> 10
<211> 632
<212> PRT
<213> Hepatitis C virus

<400> 10

```
Met Ala Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu Leu Gly
 1               5                  10                  15
Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly
            20                  25                  30
Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr Cys
        35                  40                  45
Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys Thr
    50                  55                  60
Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val Asp
65                  70                  75                  80
Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Met Thr
                85                  90                  95
Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala
            100                 105                 110
Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu
        115                 120                 125
Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Val Gly Gly Pro Leu
    130                 135                 140
Leu Cys Pro Ser Gly His Val Val Gly Ile Phe Arg Ala Ala Val Cys
145                 150                 155                 160
Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser Met
                165                 170                 175
Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro
            180                 185                 190
Ala Val Pro Gln Thr Phe Gln Val Ala His Leu His Ala Pro Thr Gly
        195                 200                 205
Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr
    210                 215                 220
Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly
225                 230                 235                 240
Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly
                245                 250                 255
Val Arg Thr Ile Thr Thr Gly Ala Pro Ile Thr Tyr Ser Thr Tyr Gly
            260                 265                 270
Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile
        275                 280                 285
Ile Cys Gln Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu Gly Ile
    290                 295                 300
Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val
305                 310                 315                 320
Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn
                325                 330                 335
Ile Glu Glu Val Ala Leu Ser Asn Asn Gly Glu Ile Pro Phe Tyr Gly
            340                 345                 350
Lys Ala Ile Pro Ile Glu Ala Ile Lys Gly Gly Arg His Leu Ile Phe
        355                 360                 365
Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Ser Gly
    370                 375                 380
Leu Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val
385                 390                 395                 400
Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
```

405 410 415
Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys
420 425 430
Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu
435 440 445
Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly
450 455 460
Arg Thr Gly Arg Gly Arg Ser Gly Ile Tyr Arg Phe Val Thr Pro Gly
465 470 475 480
Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr
485 490 495
Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser Val
500 505 510
Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
515 520 525
His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile Asp
530 535 540
Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro Tyr
545 550 555 560
Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro
565 570 575
Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr
580 585 590
Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn
595 600 605
Glu Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys Met
610 615 620
Ser Ala Asp Leu Glu Val Val Thr
625 630

<210> 11
<211> 214
<212> PRT
<213> Hepatitis C virus

<400> 11
Met Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr
1 5 10 15
Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu
20 25 30
Met Ala Phe Thr Ala Ser Ile Thr Ser Pro Leu Thr Thr Gln Asn Thr
35 40 45
Leu Leu Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Pro
50 55 60
Pro Ser Ala Ala Ser Ala Phe Val Gly Ala Gly Ile Ala Gly Ala Ala
65 70 75 80
Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly
85 90 95
Tyr Gly Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Val Met Ser
100 105 110
Gly Glu Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile
115 120 125
Leu Ser Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu
130 135 140
Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg
145 150 155 160

Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr
165 170 175
Val Pro Glu Ser Asp Ala Ala Ala Arg Val Thr Gln Ile Leu Ser Ser
180 185 190
Leu Thr Ile Thr Gln Leu Leu Lys Arg Leu His Gln Trp Ile Asn Glu
195 200 205
Asp Cys Ser Thr Pro Cys
210

<210> 12
<211> 592
<212> PRT
<213> Hepatitis C virus

<400> 12
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
1 5 10 15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Pro Leu Ser Asn Ser Leu Leu
20 25 30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Leu
35 40 45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
50 55 60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys
65 70 75 80
Ala Lys Leu Leu Ser Ile Glu Glu Ala Cys Lys Leu Thr Pro Pro His
85 90 95
Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
100 105 110
Ser Ser Arg Ala Val Asn His Ile Arg Ser Val Trp Glu Asp Leu Leu
115 120 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Ser Glu
130 135 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145 150 155 160
Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
165 170 175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
180 185 190
Gly Phe Gln Tyr Ser Pro Lys Gln Arg Val Glu Phe Leu Val Asn Thr
195 200 205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ser Tyr Gly Thr Arg Cys
210 215 220
Phe Gly Ser Thr Val Thr Glu Ser Asp Ile Arg Val Glu Glu Ser Ile
225 230 235 240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
245 250 255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
260 265 270
Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
275 280 285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Thr Ala Ala Cys
290 295 300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp
305 310 315 320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ala

```
                325                 330                 335
Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
            340                 345                 350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                 360                 365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                 375                 380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                 390                 395                 400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                 410                 415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
            420                 425                 430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile
        435                 440                 445
Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                 455                 460
Glu Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                 470                 475                 480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                 490                 495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
            500                 505                 510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Arg Tyr Leu Phe Asn
        515                 520                 525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
    530                 535                 540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                 550                 555                 560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Pro Leu Cys
                565                 570                 575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
            580                 585                 590
```

<210> 13
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 13
Pro Arg Phe Gly Lys Ala Ile Pro Ile Glu Ala Ile Lys Gly Gly
 1   5   10   15

<210> 14
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 14
Tyr Arg Leu Gly Ala Val Gln Asn Glu Val Ile Leu Thr His Pro
 1   5   10   15

<210> 15
<211> 15

<212> PRT
<213> Hepatitis C virus

<400> 15
Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
1 5 10 15

<210> 16
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 16
Ala Ala Ala Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser
1 5 10 15

<210> 17
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 17
Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala
1 5 10 15

<210> 18
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 18
Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Trp Tyr
1 5 10 15

<210> 19
<211> 9595
<212> DNA
<213> Hepatitis C virus

<400> 19
gccagccccc tgatggggc gacactccac catgaatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccccтccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacc gccgcccaca ggacgtcaag ttcccgggcg 420
gtggtcagat cgttggtgga gtttacctgt tgccgcgcag gggccccagg ttgggtgtgc 480
gcgcgactag gaaggcttcc gagcggtcgc aacctcgtgg aaggcgacaa cctatcccaa 540
aggctcgccg acccgagggc agggcctggg ctcagcccgg gtacccttgg cccctctatg 600
gcaatgaggg cctggggtgg gcaggatggc tcctgtcacc ccgcggctcc cggcctagtt 660
ggggccccac ggacccccgg cgtaggtcgc gtaacttggg taaggtcatc gataccctta 720
catgcggctt cgccgatctc atggggtaca ttccgctcgt cggcgccccc ctaggggcg 780

ctgccagggc cttggcacac ggtgtccggg ttctggagga cggcgtgaac tatgcaacag 840
ggaacttgcc cggttgctct ttctctatct tcctcttggc tctgctgtcc tgtttgacca 900
tcccagcttc cgcttatgaa gtgcgcaacg tgtccgggat ataccatgtc acgaacgact 960
gctccaactc aagcattgtg tatgaggcag cggacgtgat catgcatact cccgggtgcg 1020
tgccctgtgt tcaggagggt aacagctccc gttgctgggt agcgctcact cccacgctcg 1080
cggccaggaa tgccagcgtc cccactacga caatacgacg ccacgtcgac ttgctcgttg 1140
ggacggctgc tttctgctcc gctatgtacg tgggggatct ctgcggatct attttcctcg 1200
tctcccagct gttcaccttc tcgcctcgcc ggcatgagac agtgcaggac tgcaactgct 1260
caatctatcc cggccatgta tcaggtcacc gcatggcttg ggatatgatg atgaactggt 1320
cacctacaac agccctagtg gtgtcgcagt tgctccggat cccacaagct gtcgtggaca 1380
tggtggcggg ggcccactgg ggagtcctgg cgggccttgc ctactattcc atggtaggga 1440
actgggctaa ggttctgatt gtggcgctac tctttgccgg cgttgacggg gagacccaca 1500
cgacggggag ggtggccggc cacaccacct ccgggttcac gtccctttc tcatctgggg 1560
cgtctcagaa aatccagctt gtgaatacca acggcagctg gcacatcaac aggactgccc 1620
taaattgcaa tgactccctc caaactgggt tctttgccgc gctgttttac gcacacaagt 1680
tcaactcgtc cgggtgcccg gagcgcatgg ccagctgccg ccccattgac tggttcgccc 1740
aggggtgggg ccccatcacc tatactaagc ctaacagctc ggatcagagg ccttattgct 1800
ggcattacgc gcctcgaccg tgtggtgtcg tacccgcgtc gcaggtgtgt ggtccagtgt 1860
attgtttcac cccaagccct gttgtggtgg ggaccaccga tcgttccggt gtccctacgt 1920
atagctgggg ggagaatgag acagacgtga tgctcctcaa caacacgcgt ccgccacaag 1980
gcaactggtt cggctgtaca tggatgaata gtactgggtt cactaagacg tgcggaggtc 2040
ccccgtgtaa catcgggggg gtcggtaacc gcaccttgat ctgccccacg gactgcttcc 2100
ggaagcaccc cgaggctact tacacaaaat gtggctcggg gccctggttg acacctaggt 2160
gcctagtaga ctacccatac aggctttggc actacccctg cactctcaat ttttccatct 2220
ttaaggttag gatgtatgtg gggggcgtgg agcacaggct caatgccgca tgcaattgga 2280
ctcgaggaga gcgctgtaac ttggaggaca gggataggtc agaactcagc ccgctgctgc 2340
tgtctacaac agagtggcag atactgccct gtgctttcac caccctaccg gctttatcca 2400
ctggtttgat ccatctccat cagaacatcg tggacgtgca atacctgtac ggtgtagggt 2460
cagcgtttgt ctcctttgca atcaaatggg agtacatcct gttgcttttc cttctcctgg 2520
cagacgcgcg cgtgtgtgcc tgcttgtgga tgatgctgct gatagcccag gctgaggccg 2580
ccttagagaa cttggtggtc ctcaatgcgg cgtccgtggc cggagcgcat ggtattctct 2640
cctttcttgt gttcttctgc gccgcctggt acattaaggg caggctggct cctggggcgg 2700
cgtatgcttt ttatggcgta tggccgctgc tcctgctcct actggcgtta ccaccacgag 2760
cttacgcctt ggaccgggag atggctgcat cgtgcggggg tgcggttctt gtaggtctgg 2820
tattcttgac cttgtcacca tactacaaag tgtttctcac taggctcata tggtggttac 2880
aatactttat caccagagcc gaggcgcaca tgcaagtgtg ggtccccccc ctcaacgttc 2940
ggggaggccg cgatgccatc atcctcctca cgtgtgcggt tcatccagag ttaatttttg 3000
acatcaccaa actcctgctc gccatactcg gcccgctcat ggtgctccag gctggcataa 3060
cgagagtgcc gtacttcgtg cgcgctcaag ggctcattcg tgcatgcatg ttagtgcgaa 3120
aagtcgccgg gggtcattat gtccaaatgg tcttcatgaa gctgggcgcg ctgacaggta 3180
cgtacgttta taaccatctt accccactgc gggactgggc ccacgcgggc ctacgagacc 3240
ttgcggtggc ggtagagccc gtcgtcttct ccgccatgga gaccaaggtc atcacctggg 3300
gagcagacac cgctgcgtgt ggggacatca tcttgggtct acccgtctcc gcccgaaggg 3360
ggaaggagat atttttggga ccggctgata gtctcgaagg gcaagggtgg cgactccttg 3420
cgcccatcac ggcctactcc caacaaacgc ggggcgtact tggttgcatc atcactagcc 3480
tcacaggccg ggacaagaac caggtcgaag gggaggttca agtggtttct accgcaacac 3540
aatctttcct ggcgacctgc atcaacggcg tgtgctggac tgtctaccat ggcgctggct 3600
cgaagaccct agccggtcca aaaggtccaa tcacccaaat gtacaccaat gtagacctgg 3660
acctcgtcgg ctggcaggcg ccccccgggg cgcgctccat gacaccatgc agctgtggca 3720
gctcggacct ttacttggtc acgagacatg ctgatgtcat tccggtgcgc cggcgaggcg 3780
acagcagggg aagtctactc tcccccaggc ccgtctccta cctgaaaggc tcctcgggtg 3840
gtccattgct ttgcccttcg gggcacgtcg tgggcgtctt ccgggctgct gtgtgcaccc 3900
ggggggtcgc gaaggcggtg gacttcatac ccgttgagtc tatggaaact accatgcggt 3960
ctccggtctt cacagacaac tcaacccccc cggctgtacc gcagacattc caagtggcac 4020
atctgcacgc tcctactggc agcggcaaga gcaccaaagt gccggctgcg tatgcagccc 4080
aagggtacaa ggtgctcgtc ctgaacccgt ccgttgccgc caccttaggg tttggggcgt 4140
atatgtccaa ggcacacggt atcgacccta acatcagaac tggggtaagg accattacca 4200

cgggcggctc cattacgtac tccacctatg gcaagttcct tgccgacggt ggctgttctg 4260
ggggcgccta tgacatcata atatgtgatg agtgccactc aactgactcg actaccatct 4320
tgggcatcgg cacagtcctg gaccaagcgg agacggctgg agcgcggctc gtcgtgctcg 4380
ccaccgctac acctccggga tcggttaccg tgccacaccc caatatcgag gaaataggcc 4440
tgtccaacaa tggagagatc cccttctatg gcaaagccat ccccattgag gccatcaagg 4500
gggggaggca tctcattttc tgccattcca agaagaaatg tgacgagctc gccgcaaagc 4560
tgacaggcct cggactgaac gctgtagcat attaccgggg ccttgatgtg tccgtcatac 4620
cgcctatcgg agacgtcgtt gtcgtggcaa cagacgctct aatgacgggt ttcaccggcg 4680
attttgactc agtgatcgac tgcaatacat gtgtcaccca gacagtcgac ttcagcttgg 4740
atcccacctt caccattgag acgacgaccg tgccccaaga cgcggtgtcg cgctcgcaac 4800
ggcgaggtag aactggcagg ggtaggagtg gcatctacag gtttgtgact ccaggagaac 4860
ggccctcggg catgttcgat tcttcggtcc tgtgtgagtg ctatgacgcg ggctgtgctt 4920
ggtatgagct cacgcccgct gagacctcgg ttaggttgcg ggcttaccta aatacaccag 4980
ggttgcccgt ctgccaggac catctggagt tctgggagag cgtcttcaca ggcctcaccc 5040
acatagatgc ccacttcctg tcccagacta aacaggcagg agacaacttt ccttacctgg 5100
tggcatatca agctacagtg tgcgccaggg ctcaagctcc acctccatcg tgggaccaaa 5160
tgtggaagtg tctcatacgg ctgaaacctа cactgcacgg gccaacaccc ctgctgtata 5220
ggctaggagc cgtccaaaat gaggtcatcc tcacacaccc cataactaaa tacatcatgg 5280
catgcatgtc ggctgacctg gaggtcgtca ctagcacctg ggtgctggta ggcggagtcc 5340
ttgcagcttt ggccgcatac tgcctgacga caggcagtgt ggtcattgtg ggcaggatca 5400
tcttgtccgg gaagccagct gtcgttcccg acagggaagt cctctaccag gagttcgatg 5460
agatggaaga gtgtgcctca caacttcctt acatcgagca gggaatgcag ctcgccgagc 5520
aattcaagca aaaggcgctc gggttgttgc aaacggccac caagcaagcg gaggctgctg 5580
ctcccgtggt ggagtccaag tggcgagccc ttgagacctt ctgggcgaag cacatgtgga 5640
atttcatcag cggaatacag tacctagcag gcttatccac tctgcctgga aaccccgcga 5700
tagcatcatt gatggcattt acagcttcta tcactagccc gctcaccacc caaaacaccc 5760
tcctgtttaa catcttgggg ggatgggtgg ctgcccaact cgctcctccc agcgctgcgt 5820
cagctttcgt gggcgccggc atcgccggag cggctgttgg cagcataggc cttgggaagg 5880
tgctcgtgga catcttggcg ggctatgggg caggggtagc cggcgcactc gtggccttta 5940
aggtcatgag cggcgaggtg ccctccaccg aggacctggt caacttactc cctgccatcc 6000
tctctcctgg tgccctggtc gtcggggtcg tgtgcgcagc aatactgcgt cggcacgtgg 6060
gcccgggaga gggggctgtg cagtggatga accggctgat agcgttcgct tcgcggggta 6120
accacgtctc ccctacgcac tatgtgcctg agagcgacgc tgcagcacgt gtcactcaga 6180
tcctctctag ccttaccatc actcaactgc tgaagcggct ccaccagtgg attaatgagg 6240
actgctctac gccatgctcc ggctcgtggc taagggatgt ttgggattgg atatgcacgg 6300
tgttgactga cttcaagacc tggctccagt ccaaactcct gccgcggtta ccgggagtcc 6360
ctttcctgtc atgccaacgc gggtacaagg gagtctggcg gggggacggc atcatgcaaa 6420
ccacctgccc atgcggagca cagatcgccg gacatgtcaa aaacggttcc atgaggatcg 6480
tagggcctag aacctgcagc aacacgtggc acggaacgtt ccccatcaac gcatacacca 6540
cgggaccttg cacaccctcc ccggcgccca actattccag ggcgctatgg cgggtggctg 6600
ctgaggagta cgtggaggtt acgcgtgtgg gggatttcca ctacgtgacg ggcatgacca 6660
ctgacaacgt aaagtgccca tgccaggttc cggcccccga attcttcacg gaggtggatg 6720
gagtgcggtt gcacaggtac gctccggcgt gcaaacctct tctacgggag gacgtcacgt 6780
tccaggtcgg gctcaaccaa tacttggtcg ggtcgcagct cccatgcgag cccgaaccgg 6840
acgtaacagt gcttacttcc atgctcaccg atccctccca cattacagca gagacggcta 6900
agcgtaggct ggctagaggg tctcccccct ctttagccag ctcatcagct agccagttgt 6960
ctgcgccttc tttgaaggcg acatgcacta cccaccatga ctccccggac gctgacctca 7020
tcgaggccaa cctcttgtgg cggcaggaga tgggcggaaa catcactcgc gtggagtcag 7080
agaataaggt agtaattctg gactctttcg aaccgcttca cgcggagggg gatgagaggg 7140
agatatccgt cgcggcggag atcctgcgaa aatccaggaa gttcccctca gcgttgccca 7200
tatgggcacg cccggactac aatcctccac tgctagagtc ctggaaggac ccggactacg 7260
tccctccggt ggtacacgga tgcccattgc cacctaccaa ggctcctcca ataccacctc 7320
cacggagaaa gaggacggtt gtcctgacag aatccaatgt gtcttctgcc ttggcggagc 7380
tcgccactaa gaccttcggt agctccggat cgtcggccgt tgatagcggc acggcgaccg 7440
cccttcctga cctggcctcc gacgacggtg acaaaggatc cgacgttgag tcgtactcct 7500
ccatgccccc ccttgaaggg gagccggggg accccgatct cagcgacggg tcttggtcta 7560
ccgtgagtga ggaggctagt gaggatgtcg tctgctgctc aatgtcctat acgtggacag 7620

gcgccctgat cacgccatgc gctgcggagg aaagtaagct gcccatcaac ccgttgagca 7680
actctttgct gcgtcaccac aacatggtct acgccacaac atcccgcagc gcaagcctcc 7740
ggcagaagaa ggtcaccttt gacagattgc aagtcctgga tgatcattac cgggacgtac 7800
tcaaggagat gaaggcgaag gcgtccacag ttaaggctaa gcttctatct atagaggagg 7860
cctgcaagct gacgccccca cattcggcca aatccaaatt tggctatggg gcaaaggacg 7920
tccggaacct atccagcagg gccgttaacc acatccgctc cgtgtgggag gacttgctgg 7980
aagacactga aacaccaatt gacaccacca tcatggcaaa aagtgaggtt ttctgcgtcc 8040
aaccagagaa gggaggccgc aagccagctc gccttatcgt attcccagac ctgggagttc 8100
gtgtatgcga gaagatggcc ctttacgacg tggtctccac ccttcctcag gccgtgatgg 8160
gctcctcata cggatttcaa tactccccca agcagcgggt cgagttcctg gtgaatacct 8220
ggaaatcaaa gaaatgccct atgggcttct catatgacac ccgctgtttt gactcaacgg 8280
tcactgagag tgacattcgt gttgaggagt caatttacca atgttgtgac ttggcccccg 8340
aggccagaca ggccataagg tcgctcacag agcggcttta catcgggggt cccctgacta 8400
actcaaaagg gcagaactgc ggttatcgcc ggtgccgcgc aagtggcgtg ctgacgacta 8460
gctgcggtaa taccctcaca tgttacttga aggccactgc agcctgtcga gctgcaaagc 8520
tccaggactg cacgatgctc gtgaacggag acgaccttgt cgttatctgt gaaagcgcgg 8580
gaacccagga ggatgcggcg gccctacgag ccttcacgga ggctatgact aggtattccg 8640
ccccccccgg ggatccgccc caaccagaat acgacctgga gctgataaca tcatgttcct 8700
ccaatgtgtc agtcgcgcac gatgcatctg gcaaaagggt atactacctc acccgtgacc 8760
ccaccacccc ccttgcacgg gctgcgtggg agacagctag acacactcca atcaactctt 8820
ggctaggcaa tatcatcatg tatgcgccca ccctatgggc aaggatgatt ctgatgactc 8880
acttttctc catccttcta gctcaagagc aacttgaaaa agccctggat tgtcagatct 8940
acggggcttg ctactccatt gagccacttg acctacctca gatcattgaa cgactccatg 9000
gtcttagcgc atttacactc cacagttact ctccaggtga gatcaatagg gtggcttcat 9060
gcctcaggaa acttggggta ccacccttgc gaacctggag acatcgggcc agaagtgtcc 9120
gcgctaagct actgtcccag gggggaggg ccgccacttg tggcagatac ctctttaact 9180
gggcagtaag gaccaagctt aaactcactc caatcccggc cgcgtcccag ctggacttgt 9240
ctggctggtt cgtcgctggt tacagcgggg gagacatata tcacagcctg tctcgtgccc 9300
gaccccgctg gtttccgttg tgcctactcc tactttctgt aggggtaggc atttacctgc 9360
tccccaaccg atgaacgggg agctaaccac tccaggcctt aagccatttc ctgtttttt 9420
ttttttttt ttttttttt tctttttttt tttctttcct ttccttcttt ttttccttc 9480
ttttttcctt ctttaatggt ggctccatct tagccctagt cacggctagc tgtgaaaggt 9540
ccgtgagccg catgactgca gagagtgctg atactggcct ctctgcagat catgt 9595

<210> 20
<211> 576
<212> DNA
<213> Hepatitis C virus

<400> 20
atgagcacca accccaagcc ccagcgcaag accaagcgga acaccaaccg gagaccccag 60
gacgtcaagt tcccaggagg aggccagatc gtgggcggcg tgtacctgct gccccgccgg 120
gggccccggc tgggcgtgcg cgccacccgc aagaccagcg agcgctccca gccaagaggc 180
agacgccagc cgatcccgaa ggcccgccgc cctgagggcc gggcttgggc ccagccaggc 240
tacccctggc ccctgtatgg caacgagggc ctgggatggg ctgggtggct cctcagcccc 300
cgggggtcta ggcccagttg gggaccgacc gacccccgca ggcgcagccg caacctggga 360
aaggtgatcg acacgctcac ctgcggcttc gccgacttga tgggatacat ccctctggtg 420
ggggcccctc tgggcggagc cgcgcgcgcc ctggctcacg gggtccgggt gctcgaggac 480
ggggtgaact acgccaccgg gaacctgccc ggctgcagct tctccatctt cctgctggcg 540
ctgctgagct gcctcaccat ccccgctagc gcatga 576

<210> 21
<211> 1899
<212> DNA
<213> Hepatitis C virus

<400> 21

atggccccca tcaccgccta cagccagcag acccggggac tgctcggctg catcatcacc 60
tctctgacag gccgggataa gaaccaggtg gagggcgagg tgcaggtcgt ctcgaccgct 120
acccaaagct tcctggccac ctgtatcaac ggagtctgct ggacggtgta ccatggcgcc 180
ggcagcaaga ccctcgccgg gcctaagggc cccatcaccc agatgtacac caacgtggac 240
caggacctgg tgggctggca ggcgcccccc ggggcgagga gtatgacccc atgcacctgc 300
gggagctctg acctgtatct ggtgaccaga catgccgatg tcatcccggt gaggcgtcgc 360
ggggacagta gagggagcct gctgagcccc cgccccgtca gctacctgaa ggggtccgtg 420
ggcggccccc tgctgtgccc ctctggccac gtggtcggca tcttcagggc cgccgtgtgc 480
acgcgcggcg tggccaaggc cgtggacttt atccccgtgg agagcatgga gaccaccatg 540
cgctcccccg tgttcaccga caacagcagc ccccccgccg tgcctcagac cttccaggtc 600
gcccacctcc atgctccgac gggctccggg aagtccacga aggtgcccgc cgcgtacgcg 660
gcccagggat acaaggtgct ggtcctcaac cctagcgtgg ctgccacact cgggtttgga 720
gcgtacatga gcaaggcgca cggcatcgac cccaacatca gaactggcgt ccggaccatc 780
acaaccggcg ctcccatcac ttactctacc tacggcaagt tcctggctga tgggggtgt 840
agtgggggcg cgtacgatat tatcatctgc caggagtgcc actctaccga cagcaccaca 900
atcctgggca tcggcaccgt cctcgaccag gctgagacag cgggcgcccg cctggtggtg 960
ctggccacgg ccactccccc cggctccgtc acggtgcccc accccaatat cgaggaggtg 1020
gccctgagca acaacggcga gatcccattc tacggcaagg ctatcccgat cgaggcgatt 1080
aagggaggca gacatctgat cttctgccac agcaagaaga agtgcgacga gctcgccgcc 1140
aagctgagcg gcctcggact caacgccgtg gcttactaca ggggactgga cgtgtccgtg 1200
atcccgacca gcggagacgt ggtggtcgtg gccaccgacg ccctgatgac cggcttcacc 1260
ggagacttcg acagcgtcat cgactgcaac acctgcgtga cccagaccgt ggacttcagc 1320
ctggacccca ccttcaccat cgagaccacc acagtgcccc aggacgccgt gtcccgcagc 1380
cagcgccggg gccggaccgg ccgcggccgg agtggcatct ataggttcgt gaccccgggc 1440
gagcgcccca gcggcatgtt cgatagttcc gtgctgtgcg agtgctacga cgccggatgc 1500
gcgtggtacg agctgacccc ggcggagacc tctgtccgcc tgagggctta cttgaatacc 1560
ccgggcctgc ccgtgtgcca ggatcatctc gagttctggg aatccgtctt caccggcctg 1620
acacacatcg acgcccattt cttgtcccaa accaagcagg ctggcgacaa tttcccgtat 1680
ctggtcgcgt accaggccac ggtgtgcgcg cgtgcgcagg ctccccccc tagctgggat 1740
cagatgtgga agtgcctgat ccgcctgaag cccaccctgc atgggcccac ccccctgctg 1800
taccgcctgg gcgcggtgca gaacgaagtc accttgaccc accccatcac caagtacatc 1860
atggcgtgca tgtccgctga cctggaggtg gtcacctga 1899

<210> 22
<211> 645
<212> DNA
<213> Hepatitis C virus

<400> 22
atgttttggg ccaagcatat gtggaacttc atcagcggca tccagtacct cgccgggctg 60
agcaccctcc cgggcaaccc cgcgatcgca agcctgatgg cgttcacagc gagcatcacc 120
tccccctga ctacccagaa cacactgctg ttcaacatcc tgggggctg ggtcgccgct 180
cagctggccc ctccttccgc cgccagcgcc tttgtggggg cgggaatcgc cggggccgcc 240
gtcggctcca tcggactggg caaggtgctg gtcgacatcc tggcgggcta cggcgcggga 300
gtcgccggag ccctggtggc cttcaaggtg atgagcggag aggtgccaag cactgaggac 360
ctggtgaacc tgctgccggc gatcctgagc ccgggcgccc tggtggtggg cgtggtgtgt 420
gctgccatcc tcaggcgcca cgtgggcccg ggcgagggag ccgtgcagtg gatgaaccgc 480
ctgatcgcct ttgcctcccg cggcaaccac gtcagcccta cacattacgt gcccgagagc 540
gatgccgccg cccgcgtgac ccagatcctg agctccctga ccatcaccca gctgctcaag 600
aggctgcacc agtggatcaa cgaggactgc tccacccctt gctga 645

<210> 23
<211> 1779
<212> DNA
<213> Hepatitis C virus

<400> 23

```
atgtccatgt cctacacctg gaccggcgcc ctgatcaccc cctgcgccgc cgaggagagc 60
aagctcccga ttaacccccc gtccaactct ctgctccgcc atcacaacat ggtgtatgcc 120
accacctccc gctctgcgag cctccgccag aagaaggtga cgttcgacag actgcaggtg 180
ctggacgacc attacaggga cgtgctgaag gaaatgaagg ccaaggctag caccgtgaag 240
gccaagctgc tcagcattga ggaggcttgc aagctgaccc ccccccacag tgctaaatcc 300
aagttcggct acggcgccaa ggacgtgagg aacctgtcct cgcgcgctgt gaaccatatc 360
cgcagcgtgt gggaggacct gctcgaggac accgagaccc ccatcgacac aaccatcatg 420
gccaagtccg aggtgttctg cgtgcagccg gagaaaggag gccgcaagcc agcccgcctg 480
atcgtcttcc ccgacctggg cgtgagagtc tgcgagaaga tggccctcta cgacgtggtg 540
tccaccctgc cgcaggccgt gatggggagt tcctacggct tccagtacag cccgaagcag 600
agggtggagt tcctggtgaa cacgtggaag tctaagaaat gccccatggg gttcagttac 660
ggaacaaggt gcttcgggag tactgtgacc gaatccgata tccgcgtgga ggagagcatc 720
taccagtgtt gtgacctcgc ccccgaggcg agacaggcca tccgctccct gaccgagagg 780
ctgtatatcg gcggcccact gaccaacagc aaggggcaga actgcggcta tcgccgttgt 840
cgggcctccg gggtgctcac cacctcttgc gggaacaccc tcacctgcta cctcaaggcg 900
accgctgcct gcagagccgc gaagctgcag gactgcacca tgctcgtgaa cggcgacgat 960
ctggtggtga tctgtgagtc cgcgggcacg caggaggacg cggcggccct gcgggcgttc 1020
acagaggcca tgacacgcta cagtgccccc cccggcgacc ccccccagcc cgaatacgat 1080
ctggagctca tcactagttg cagctcgaac gtgtctgtgg cccatgacgc ttctggcaaa 1140
cgggtgtatt atctgacgcg cgatcccacc acccccctcg ccagagccgc gtgggagaca 1200
gctcggcaca cccctgtgaa ctcttggctg ggcaacatca tcatgtacgc ccctaccctg 1260
tgggctcgca tgatcctgat gacccacttc ttcagtatcc tcctcgctca ggagcagctg 1320
gagaaggcgc tcgactgcca gatctacggc gcctgctata gtatcgagcc tctcgacctg 1380
ccccagatca tcgagagact gcatgggctc agcgccttct ccctccatag ttactctcct 1440
ggagaaatta accgggtggc gagctgtctg cggaagctcg gcgtcccccc tctgcgcgtt 1500
tggcggcatc gcgccaggag tgtgagggcc aagctgctga gccagggcgg aagggccgcc 1560
acctgcggcc ggtatctctt caactgggcc gtgcgcacca agctcaagct caccccccatc 1620
cctgccgcca gtcagctgga tctcagtggg tggttcgtgg ccggctattc tggcggcgac 1680
atctaccact ccctcagcag ggcgcgcccc cgctggttcc ccctgtgcct gctgctcctg 1740
agcgtcggag tcggcatcta cctgctgccc aaccgctga                          1779
```

<210> 24
<211> 3010
<212> PRT
<213> Hepatitis C virus

<400> 24

```
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
 1               5                  10                  15
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25                  30
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
        35                  40                  45
Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
    50                  55                  60
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
65                  70                  75                  80
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
                85                  90                  95
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
            100                 105                 110
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
        115                 120                 125
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
    130                 135                 140
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145                 150                 155                 160
```

Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
165 170 175
Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr
180 185 190
Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys Ser
195 200 205
Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Val Ile Met His Thr Pro
210 215 220
Gly Cys Val Pro Cys Val Gln Glu Gly Asn Ser Ser Arg Cys Trp Val
225 230 235 240
Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr
245 250 255
Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Thr Ala Ala Phe Cys
260 265 270
Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Ile Phe Leu Val Ser
275 280 285
Gln Leu Phe Thr Phe Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys
290 295 300
Asn Cys Ser Ile Tyr Pro Gly His Val Ser Gly His Arg Met Ala Trp
305 310 315 320
Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln
325 330 335
Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His
340 345 350
Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp
355 360 365
Ala Lys Val Leu Ile Val Ala Leu Leu Phe Ala Gly Val Asp Gly Glu
370 375 380
Thr His Thr Thr Gly Arg Val Ala Gly His Thr Thr Ser Gly Phe Thr
385 390 395 400
Ser Leu Phe Ser Ser Gly Ala Ser Gln Lys Ile Gln Leu Val Asn Thr
405 410 415
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
420 425 430
Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Ala His Lys Phe Asn
435 440 445
Ser Ser Gly Cys Pro Glu Arg Met Ala Ser Cys Arg Pro Ile Asp Trp
450 455 460
Phe Ala Gln Gly Trp Gly Pro Ile Thr Tyr Thr Lys Pro Asn Ser Ser
465 470 475 480
Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Val
485 490 495
Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser
500 505 510
Pro Val Val Val Gly Thr Thr Asp Arg Ser Gly Val Pro Thr Tyr Ser
515 520 525
Trp Gly Glu Asn Glu Thr Asp Val Met Leu Leu Asn Asn Thr Arg Pro
530 535 540
Pro Gln Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe
545 550 555 560
Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Val Gly Asn
565 570 575
Arg Thr Leu Ile Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala
580 585 590
Thr Tyr Thr Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Leu
595 600 605
Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Leu Asn Phe

610 615 620
Ser Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Leu
625 630 635 640
Asn Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asn Leu Glu Asp
645 650 655
Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp
660 665 670
Gln Ile Leu Pro Cys Ala Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly
675 680 685
Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly
690 695 700
Val Gly Ser Ala Phe Val Ser Phe Ala Ile Lys Trp Glu Tyr Ile Leu
705 710 715 720
Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp
725 730 735
Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val
740 745 750
Val Leu Asn Ala Ala Ser Val Ala Gly Ala His Gly Ile Leu Ser Phe
755 760 765
Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Ala Pro
770 775 780
Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu
785 790 795 800
Leu Ala Leu Pro Pro Arg Ala Tyr Ala Leu Asp Arg Glu Met Ala Ala
805 810 815
Ser Cys Gly Gly Ala Val Leu Val Gly Leu Val Phe Leu Thr Leu Ser
820 825 830
Pro Tyr Tyr Lys Val Phe Leu Thr Arg Leu Ile Trp Trp Leu Gln Tyr
835 840 845
Phe Ile Thr Arg Ala Glu Ala His Met Gln Val Trp Val Pro Pro Leu
850 855 860
Asn Val Arg Gly Gly Arg Asp Ala Ile Ile Leu Leu Thr Cys Ala Val
865 870 875 880
His Pro Glu Leu Ile Phe Asp Ile Thr Lys Leu Leu Leu Ala Ile Leu
885 890 895
Gly Pro Leu Met Val Leu Gln Ala Gly Ile Thr Arg Val Pro Tyr Phe
900 905 910
Val Arg Ala Gln Gly Leu Ile Arg Ala Cys Met Leu Val Arg Lys Val
915 920 925
Ala Gly Gly His Tyr Val Gln Met Val Phe Met Lys Leu Gly Ala Leu
930 935 940
Thr Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala
945 950 955 960
His Ala Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe
965 970 975
Ser Ala Met Glu Thr Lys Val Ile Thr Trp Gly Ala Asp Thr Ala Ala
980 985 990
Cys Gly Asp Ile Ile Leu Gly Leu Pro Val Ser Ala Arg Arg Gly Lys
995 1000 1005
Glu Ile Phe Leu Gly Pro Ala Asp Ser Leu Glu Gly Gln Gly Trp Arg
1010 1015 1020
Leu Leu Ala Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Val Leu
1025 1030 1035 1040
Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu
1045 1050 1055
Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr
1060 1065 1070

Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys
1075 1080 1085
Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val
1090 1095 1100
Asp Leu Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Met
1105 1110 1115 1120
Thr Pro Cys Ser Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His
1125 1130 1135
Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu
1140 1145 1150
Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro
1155 1160 1165
Leu Leu Cys Pro Ser Gly His Val Val Gly Val Phe Arg Ala Ala Val
1170 1175 1180
Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser
1185 1190 1195 1200
Met Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Thr Pro
1205 1210 1215
Pro Ala Val Pro Gln Thr Phe Gln Val Ala His Leu His Ala Pro Thr
1220 1225 1230
Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly
1235 1240 1245
Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe
1250 1255 1260
Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr
1265 1270 1275 1280
Gly Val Arg Thr Ile Thr Thr Gly Gly Ser Ile Thr Tyr Ser Thr Tyr
1285 1290 1295
Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile
1300 1305 1310
Ile Ile Cys Asp Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu Gly
1315 1320 1325
Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val
1330 1335 1340
Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro
1345 1350 1355 1360
Asn Ile Glu Glu Ile Gly Leu Ser Asn Asn Gly Glu Ile Pro Phe Tyr
1365 1370 1375
Gly Lys Ala Ile Pro Ile Glu Ala Ile Lys Gly Gly Arg His Leu Ile
1380 1385 1390
Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Thr
1395 1400 1405
Gly Leu Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser
1410 1415 1420
Val Ile Pro Pro Ile Gly Asp Val Val Val Val Ala Thr Asp Ala Leu
1425 1430 1435 1440
Met Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr
1445 1450 1455
Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile
1460 1465 1470
Glu Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg
1475 1480 1485
Gly Arg Thr Gly Arg Gly Arg Ser Gly Ile Tyr Arg Phe Val Thr Pro
1490 1495 1500
Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys
1505 1510 1515 1520
Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser

1525 1530 1535
Val Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln
1540 1545 1550
Asp His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile
1555 1560 1565
Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro
1570 1575 1580
Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro
1585 1590 1595 1600
Pro Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro
1605 1610 1615
Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln
1620 1625 1630
Asn Glu Val Ile Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys
1635 1640 1645
Met Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly
1650 1655 1660
Gly Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Thr Thr Gly Ser Val
1665 1670 1675 1680
Val Ile Val Gly Arg Ile Ile Leu Ser Gly Lys Pro Ala Val Val Pro
1685 1690 1695
Asp Arg Glu Val Leu Tyr Gln Glu Phe Asp Glu Met Glu Glu Cys Ala
1700 1705 1710
Ser Gln Leu Pro Tyr Ile Glu Gln Gly Met Gln Leu Ala Glu Gln Phe
1715 1720 1725
Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Thr Lys Gln Ala Glu
1730 1735 1740
Ala Ala Ala Pro Val Val Glu Ser Lys Trp Arg Ala Leu Glu Thr Phe
1745 1750 1755 1760
Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala
1765 1770 1775
Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala
1780 1785 1790
Phe Thr Ala Ser Ile Thr Ser Pro Leu Thr Thr Gln Asn Thr Leu Leu
1795 1800 1805
Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Pro Pro Ser
1810 1815 1820
Ala Ala Ser Ala Phe Val Gly Ala Gly Ile Ala Gly Ala Ala Val Gly
1825 1830 1835 1840
Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly
1845 1850 1855
Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Val Met Ser Gly Glu
1860 1865 1870
Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser
1875 1880 1885
Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg
1890 1895 1900
His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile
1905 1910 1915 1920
Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro
1925 1930 1935
Glu Ser Asp Ala Ala Ala Arg Val Thr Gln Ile Leu Ser Ser Leu Thr
1940 1945 1950
Ile Thr Gln Leu Leu Lys Arg Leu His Gln Trp Ile Asn Glu Asp Cys
1955 1960 1965
Ser Thr Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Ile
1970 1975 1980

Cys Thr Val Leu Thr Asp Phe Lys Thr Trp Leu Gln Ser Lys Leu Leu
1985 1990 1995 2000
Pro Arg Leu Pro Gly Val Pro Phe Leu Ser Cys Gln Arg Gly Tyr Lys
2005 2010 2015
Gly Val Trp Arg Gly Asp Gly Ile Met Gln Thr Thr Cys Pro Cys Gly
2020 2025 2030
Ala Gln Ile Ala Gly His Val Lys Asn Gly Ser Met Arg Ile Val Gly
2035 2040 2045
Pro Arg Thr Cys Ser Asn Thr Trp His Gly Thr Phe Pro Ile Asn Ala
2050 2055 2060
Tyr Thr Thr Gly Pro Cys Thr Pro Ser Pro Ala Pro Asn Tyr Ser Arg
2065 2070 2075 2080
Ala Leu Trp Arg Val Ala Ala Glu Glu Tyr Val Glu Val Thr Arg Val
2085 2090 2095
Gly Asp Phe His Tyr Val Thr Gly Met Thr Thr Asp Asn Val Lys Cys
2100 2105 2110
Pro Cys Gln Val Pro Ala Pro Glu Phe Phe Thr Glu Val Asp Gly Val
2115 2120 2125
Arg Leu His Arg Tyr Ala Pro Ala Cys Lys Pro Leu Leu Arg Glu Asp
2130 2135 2140
Val Thr Phe Gln Val Gly Leu Asn Gln Tyr Leu Val Gly Ser Gln Leu
2145 2150 2155 2160
Pro Cys Glu Pro Glu Pro Asp Val Thr Val Leu Thr Ser Met Leu Thr
2165 2170 2175
Asp Pro Ser His Ile Thr Ala Glu Thr Ala Lys Arg Arg Leu Ala Arg
2180 2185 2190
Gly Ser Pro Pro Ser Leu Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala
2195 2200 2205
Pro Ser Leu Lys Ala Thr Cys Thr Thr His His Asp Ser Pro Asp Ala
2210 2215 2220
Asp Leu Ile Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn
2225 2230 2235 2240
Ile Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe
2245 2250 2255
Glu Pro Leu His Ala Glu Gly Asp Glu Arg Glu Ile Ser Val Ala Ala
2260 2265 2270
Glu Ile Leu Arg Lys Ser Arg Lys Phe Pro Ser Ala Leu Pro Ile Trp
2275 2280 2285
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Leu Glu Ser Trp Lys Asp Pro
2290 2295 2300
Asp Tyr Val Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Thr Lys
2305 2310 2315 2320
Ala Pro Pro Ile Pro Pro Pro Arg Arg Lys Arg Thr Val Val Leu Thr
2325 2330 2335
Glu Ser Asn Val Ser Ser Ala Leu Ala Glu Leu Ala Thr Lys Thr Phe
2340 2345 2350
Gly Ser Ser Gly Ser Ser Ala Val Asp Ser Gly Thr Ala Thr Ala Leu
2355 2360 2365
Pro Asp Leu Ala Ser Asp Asp Gly Asp Lys Gly Ser Asp Val Glu Ser
2370 2375 2380
Tyr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu
2385 2390 2395 2400
Ser Asp Gly Ser Trp Ser Thr Val Ser Glu Glu Ala Ser Glu Asp Val
2405 2410 2415
Val Cys Cys Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro
2420 2425 2430
Cys Ala Ala Glu Glu Ser Lys Leu Pro Ile Asn Pro Leu Ser Asn Ser

2435 2440 2445
Leu Leu Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala
2450 2455 2460
Ser Leu Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp
2465 2470 2475 2480
Asp His Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr
2485 2490 2495
Val Lys Ala Lys Leu Leu Ser Ile Glu Glu Ala Cys Lys Leu Thr Pro
2500 2505 2510
Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg
2515 2520 2525
Asn Leu Ser Ser Arg Ala Val Asn His Ile Arg Ser Val Trp Glu Asp
2530 2535 2540
Leu Leu Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys
2545 2550 2555 2560
Ser Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala
2565 2570 2575
Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met
2580 2585 2590
Ala Leu Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser
2595 2600 2605
Ser Tyr Gly Phe Gln Tyr Ser Pro Lys Gln Arg Val Glu Phe Leu Val
2610 2615 2620
Asn Thr Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ser Tyr Asp Thr
2625 2630 2635 2640
Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Val Glu Glu
2645 2650 2655
Ser Ile Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile
2660 2665 2670
Arg Ser Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser
2675 2680 2685
Lys Gly Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu
2690 2695 2700
Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Thr Ala
2705 2710 2715 2720
Ala Cys Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly
2725 2730 2735
Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala
2740 2745 2750
Ala Ala Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro
2755 2760 2765
Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser
2770 2775 2780
Cys Ser Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val
2785 2790 2795 2800
Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp
2805 2810 2815
Glu Thr Ala Arg His Thr Pro Ile Asn Ser Trp Leu Gly Asn Ile Ile
2820 2825 2830
Met Tyr Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe
2835 2840 2845
Phe Ser Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys
2850 2855 2860
Gln Ile Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln
2865 2870 2875 2880
Ile Ile Glu Arg Leu His Gly Leu Ser Ala Phe Thr Leu His Ser Tyr
2885 2890 2895

Ser Pro Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly
2900 2905 2910
Val Pro Pro Leu Arg Thr Trp Arg His Arg Ala Arg Ser Val Arg Ala
2915 2920 2925
Lys Leu Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Arg Tyr Leu
2930 2935 2940
Phe Asn Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala
2945 2950 2955 2960
Ala Ser Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly
2965 2970 2975
Gly Asp Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Pro
2980 2985 2990
Leu Cys Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro
2995 3000 3005
Asn Arg
3010